

Выводы

Предложен и реализован новый алгоритм для распознавания руки на основе *SURF*-дескрипторов и нейронной сети. Изложен новый метод генерации дескрипторов для нейронной сети, которые устойчивы к поворотам и масштабированию, частично к изменению освещения, и робастны на сложном фоне. Созданные дескрипторы позволяют распознавать левую и правую руки одновременно, используя только один классификатор, при этом обучение требуется только один раз для одной руки. Алгоритм

позволяет распознавать форму руки с высокой точностью. На основе предложенного алгоритма реализовано программное обеспечение для распознавания формы руки в видеопотоке в режиме реального времени. Программа поддерживает генерацию «словаря» признаков, формирование базы данных для обучения, создание и обучение нейронной сети. Проведенные численные эксперименты по распознаванию формы руки на видеопоследовательности в реальном времени показали, что средняя точность распознавания составляет 92 %.

СПИСОК ЛИТЕРАТУРЫ

1. Zhou H., Huang T.S. Tracking articulated hand motion with Eigen dynamics analysis // Proc. of Intern. Conf. on Computer Vision. – Nice (France), 2003. – V. 2. – P. 1102–1109.
2. Chen F., Fu C., Huang C. Hand gesture recognition using a real-time tracking method and Hidden Markov Models // Image and Vision Computing. – Oxford: Elsevier, 2003. – V. 21. – № 8. – P. 745–758.
3. Ng C.W., Ranganath S. Gesture recognition via pose classification // Proc. 15th Intern. Conf. on Pattern Recognition. – Barcelona (Spain), 2000. – V. 3. – P. 699–704.
4. Oka K., Sato Y., Koike H. Real-time fingertip tracking and gesture recognition // Proc. IEEE Computer Graphics and Applications. – 2002. – V. 22. – № 6. – P. 64–71.
5. Huang C., Jeng S. A model-based hand gesture recognition system // Machine Vision and Application. – 2001. – V. 12. – № 5. – P. 243–258.
6. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features // IEEE CVPR01. – 2001. – P. 511–518.
7. Kolsch M., Turk M. Robust hand detection // Proc. of the 6th IEEE Intern. Conf. on Automatic Face and Gesture Recognition. – Seoul (Korea), May 2004. – Seoul, 2004. – P. 614–619.
8. Alper Y., Omar J., Mubarak S. Object tracking: A Survey // ACM Computing Surveys (CSUR). – 2006. – V. 38. – Iss. 4. – Art. № 13.
9. Bradski G.R. Real time face and object tracking as a component of a perceptual user interface // Proc. of the 4th IEEE Workshop on Applications of Computer Vision. – New Jersey (USA), October 1998. – New Jersey, 1998. – P. 214–219.
10. Bay H., Ess A., Tuytelaars T., Van Gool L. SURF: Speeded Up Robust Features // Computer Vision and Image Understanding. – Oxford: Elsevier, 2008. – V. 110. – № 3. – P. 346–359.

Поступила 11.03.2012 г.

УДК 004.932

РАСПОЗНАВАНИЕ ЛИЦ НА ОСНОВЕ ПРИМЕНЕНИЯ МЕТОДА ВИОЛЫ–ДЖОНСА, ВЕЙВЛЕТ–ПРЕОБРАЗОВАНИЯ И МЕТОДА ГЛАВНЫХ КОМПОНЕНТ

Буй Тхи Тху Чанг, Фан Нгок Хоанг, В.Г. Спицын

Томский политехнический университет
E-mail: trangbt.084@gmail.com

На основе совместного применения метода Виолы–Джонса, вейвлет-преобразования и метода главных компонент предложен новый алгоритм распознавания лиц на цифровых изображениях и видеопоследовательностях в режиме реального времени. Описан алгоритм и разработано программное обеспечение для распознавания лиц. Приведен пример работы и представлены результаты тестирования программы. Показано, что использование предложенного оригинального алгоритма дает возможность эффективного распознавания лиц на цифровых изображениях и видеопоследовательностях.

Ключевые слова:

Распознавание лиц, метод Виолы–Джонса, метод главных компонент, вейвлет-преобразование.

Key words:

Face recognition, method Viola-Jones, principal component analysis, wavelet transform.

Введение

Распознавание лиц является одной из самых изученных задач в таких областях как цифровая обработка изображений, компьютерное зрение, биометрия, организация видеоконференций, создание интеллектуальных систем безопасности и контроля доступа и т. п. Процесс распознавания лиц

обычно состоит из двух этапов: поиск области лица на изображении, и сравнение найденного лица с лицами, находящимися в базе данных. В настоящее время метод Виолы–Джонса [1] является самым популярным методом для поиска области лица на изображении из-за его высокой скорости и эффективности. Детектор лица Виолы–Джонса

основан на основных идеях: интегральном представлении изображения, методе построения классификатора на основе алгоритма адаптивного бустинга (*AdaBoost*), и методе комбинирования классификаторов в каскадную структуру. Эти идеи позволяют построить детектор лица, способный работать в режиме реального времени. Метод главных компонент и вейвлет-преобразование используются для получения характеристик изображения. В задаче распознавания лиц они успешно применяются для сравнения компонент, характеризующих цветные изображения, с компонентами, описывающими неизвестные изображения.

Целью данной работы является создание нового алгоритма, основанного на комбинации метода Виолы–Джонса, вейвлет-преобразования и метода главных компонент (МГК) для распознавания лиц на цифровых изображениях и видеопоследовательностях в режиме реального времени.

1. Метод Виолы–Джонса

Метод был разработан и представлен в 2001 г. Полом Виолой и Майклом Джонсом. Он до сих пор является эффективным методом для поиска объектов на изображениях и видеопоследовательностях в режиме реального времени [1, 2]. Следует отметить, что этот детектор обладает крайне низкой вероятностью ложного обнаружения лица. Метод хорошо работает и обнаруживает черты лица даже при наблюдении объекта под небольшим углом, примерно до 30°. При угле наклона больше 30° вероятность обнаружения лица резко падает. Указанная особенность метода не позволяет в стандартной реализации детектировать лицо человека, повернутое под произвольным углом, что в значительной мере затрудняет или делает невозможным использование алгоритма в современных производственных системах с учетом их растущих потребностей.

1.1. Интегральное представление изображений

Для того, чтобы рассчитать яркость прямоугольного участка изображения, используют интегральное представление [3]. Такое представление используется часто и в других методах, например, в вейвлет-преобразованиях, *Speeded up robust feature* (SURF), фильтрах Хаара и многих разработанных алгоритмах. Интегральное представление позволяет быстро рассчитывать суммарную яркость произвольного прямоугольника на данном изображении, причем время расчета не зависит от площади прямоугольника.

Интегральное представление изображения представляет собой матрицу, совпадающую по размерам с исходным изображением. В каждом ее элементе хранится сумма интенсивностей всех пикселей, находящихся левее и выше данного элемента. Элементы матрицы рассчитываются по следующей формуле:

$$I(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'),$$

где $I(x, y)$ – значение точки (x, y) интегрального изображения; $i(x, y)$ – значение интенсивности исходного изображения. На основе применения интегрального представления изображения вычисление признаков одинакового вида, но с разными геометрическими параметрами, происходит за одинаковое время.

Каждый элемент матрицы $I(x, y)$ представляет собой сумму пикселей в прямоугольнике от $i(0, 0)$ до $i(x, y)$, т. е. значение каждого элемента $I(x, y)$ равно сумме значений всех пикселей левее и выше данного пикселя $i(x, y)$. Расчет матрицы занимает линейное время, пропорциональное числу пикселей в изображении и его можно производить по следующей формуле:

$$I(x, y) = i(x, y) - I(x-1, y-1) + \\ + I(x, y-1) + I(x-1, y).$$

Интегральное представление имеет интересную особенность. По интегральной матрице можно очень быстро вычислить сумму пикселей произвольного прямоугольника.

1.2. Хаар-подобные характеристики

С точки зрения необходимости использования достаточно простых алгоритмов получения признаков, перспективным является использование Хаар-подобных характеристик, представляющих собой результат сравнения яркостей в двух прямоугольных областях изображения.

Предположим, что задано множество объектов A и множество допустимых ответов B . Пусть $g: A \rightarrow B$ называется решающей функцией. Решающая функция g должна допускать эффективную компьютерную реализацию, по этой причине её также называют алгоритмом. Признак (*feature*) f объекта a – отображение $f: A \rightarrow D_f$, где D_f – множество допустимых значений признака. В частности, любой алгоритм $g: A \rightarrow B$ также можно рассматривать как признак. Если задан набор признаков f_1, \dots, f_n , то вектор $x = (f_1(a), \dots, f_n(a))$ называется признаковым описанием объекта $a \in A$. Признаковые описания допустимо отождествлять с самими объектами. При этом множество $A = D_{f_1} \times \dots \times D_{f_n}$ называют признаковым пространством [4].

В стандартном методе Виолы–Джонса используются прямоугольные признаки, рис. 1. Эти признаки называются примитивами Хаара.

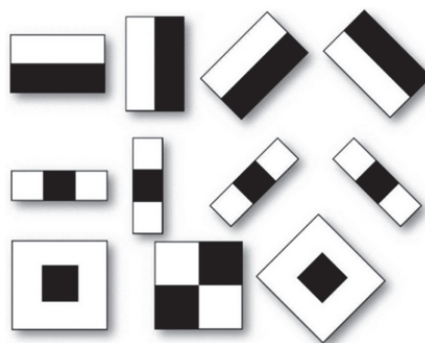


Рис. 1. Примитивы признаков Хаара

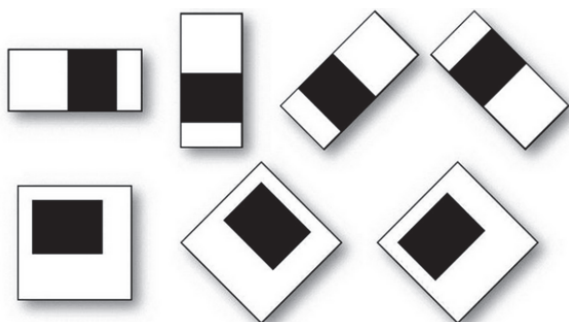


Рис. 2. Дополнительные признаки Хаара

В расширенном методе Виолы–Джонса, представленном в библиотеке OpenCV, используются дополнительные признаки (рис. 2).

Вычисляемым значением такого признака будет:

$$F = U - V,$$

где U – сумма значений яркостей точек, закрываемых светлой частью признака, а V – сумма значений яркостей точек, закрываемых темной частью признака. Для их вычисления используется понятие интегрального изображения. Хаар-подобные признаки описывают значение перепада яркости по оси X и Y изображения соответственно.

1.3. Метод построения классификатора на основе алгоритма бустинга

Бустинг – комплекс методов, способствующих повышению точности аналитических моделей. Бустинг (*boosting*) означает дословно «усиление» «слабых» моделей [5] – это процедура последовательного построения композиции алгоритмов машинного обучения, когда каждый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов. Идея бустинга была предложена Робертом Шапиро (*Schapire*) в конце 90-х гг. прошлого века [6], когда надо было найти решение вопроса о том, каким образом имея множество плохих (незначительно отличающихся от случайных) алгоритмов обучения, получить один хороший.

В результате работы алгоритма бустинга на каждой итерации формируется простой классификатор вида:

$$h_j(z) = \begin{cases} 1, & \text{если } p_j f_j(z) < p_j \theta_j, \\ 0, & \text{иначе} \end{cases},$$

где p_j – показывает направление знака неравенства; θ_j – значение порога; $f_j(z)$ – вычисленное значение признака; z – окно изображения размером 24×24 пикселей. Полученный классификатор имеет минимальную ошибку по отношению к текущим значениям весов, задействованным в процедуре обучения для определения ошибки.

1.4. Метод комбинирования классификаторов в каскадную структуру

Каскадная структура повышает скорость обнаружения, фокусируя свою работу на наиболее информативных областях изображения.

Структура каскадного детектора приведена на рис. 3. Каскад состоит из слоев, которые представляют собой классификаторы, обученные с помощью процедуры бустинга.

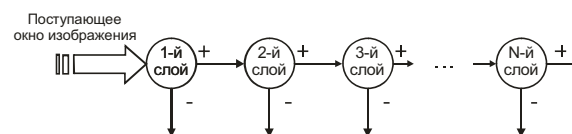


Рис. 3. Структура каскадного детектора

2. Вейвлет-преобразование

Вейвлет-преобразование широко используется для анализа нестационарных процессов. Оно показало свою эффективность для решения широкого класса задач, связанных с обработкой изображения. Коэффициенты вейвлет-преобразования содержат информацию об анализируемом процессе и используемом вейвлете. Поэтому выбор анализирующего вейвлета определяется тем, какую информацию необходимо извлечь из процесса. Каждый вейвлет имеет характерные особенности во временной и частотной областях, поэтому иногда с помощью разных вейвлетов можно полнее выявить и подчеркнуть те или иные свойства анализируемого процесса.

В работах [7, 8] представлены разложение изображения и извлечение его признаков для классификации изображений самолетов на основе применения вейвлет-преобразования Хаара и многослойной нейронной сети. В данной работе используются вейвлет-преобразования Хаара и Добеши для извлечения признаков изображения лиц. Примеры применения вейвлет-преобразования Хаара и вейвлет-преобразования Добеши для извлечения признаков изображения лица представлены на рис. 4.

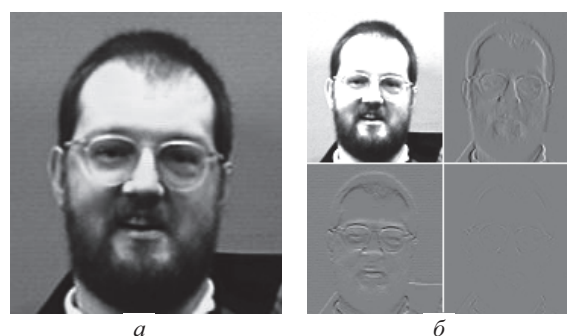


Рис. 4. Пример извлечения признаков лица: а) исходное изображение лица; б) результат после применения вейвлет-преобразования Хаара

3. Метод главных компонент

Метод главных компонент (*Principal Component Analysis, PCA*) – один из наиболее распространенных методов для уменьшения размерности данных, потери наименьшего количества информации. Он заключается в линейном ортогональном преобразовании входного вектора P размерности N в вы-

ходной вектор Q размерности M , $M < N$. Компоненты вектора Q являются некоррелированными, и общая дисперсия после преобразования остаётся неизменной.

Главная идея МГК состоит в представлении изображений лиц людей в виде набора главных компонент изображений, называемых «собственные лица» (*Eigenfaces*). Собственные лица имеют полезное свойство, заключающееся в том, что изображение, соответствующее каждому такому вектору имеет лицеподобную форму, рис. 5. Вычисление главных компонент сводится к вычислению собственных векторов и собственных значений ковариационной матрицы, которая рассчитывается из изображения. Сумма главных компонент, умноженных на соответствующие собственные вектора, является реконструкцией изображения, рис. 6.



Рис. 5. Пример изображений собственных векторов (собственные лица)

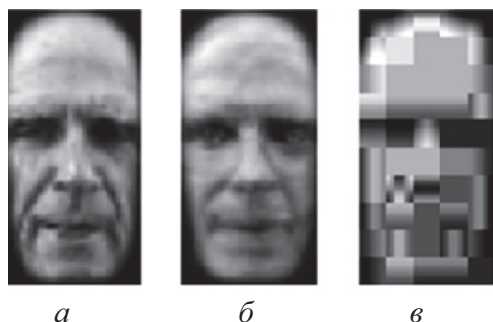


Рис. 6. Пример изображений лица: а) выровненное изображение; б) реконструкция по 85-и главным компонентам; в) JPEG-реконструкция (530 байт)

Для каждого изображения лица вычисляются его главные компоненты. Обычно берётся от 5 до 200 главных компонент. Остальные компоненты кодируют мелкие различия между лицами и шумами. Процесс распознавания заключается в сравнении главных компонент неизвестного изображения с компонентами всех известных изображений. При этом предполагается, что изображения лиц, соответствующих одному человеку, сгруппированы в кластеры в собственном пространстве. Из базы данных выбираются изображения-кандидаты, имеющие наименьшее расстояние от входного (неизвестного) изображения [9].

4. Алгоритм распознавания лиц

Целью данной работы является распознавание лиц с высокой скоростью. Для решения задачи распознавания лиц предложен алгоритм на основе применения метода Виолы–Джонса, вейвлет-преобразования и метода главных компонент. Предло-

женный алгоритм состоит из двух процессов: сохранения признаков известных лиц в базе данных и распознавания лиц.

Процесс сохранения признаков известных лиц происходит следующим образом:

- Шаг 1.** Преобразование изображения видео фрейма в полутоновое изображение.
- Шаг 2.** Применение к полутоновому изображению метода Виолы–Джонса для поиска области лица.
- Шаг 3.** Уменьшение размера области лица до 64×64 пикселей.
- Шаг 4.** Применение к полученному на шаге 3 изображению вейвлет-преобразования для извлечения признаков лица (вейвлет-коэффициентов).
- Шаг 5.** Сохранение извлеченных признаков в базе данных.

В процессе распознавания неизвестного лица осуществляются шаги 1–4, затем на основе применения метода главных компонент происходит сокращение числа признаков и их сравнение с признаками, хранящимися в базе данных. Функциональная схема предложенного алгоритма представлена на рис. 7.

5. Эксперименты

Для распознавания лиц на изображениях и видеопоследовательностях в реальном времени на языке объектно-ориентированного программирования C# (*Visual studio 2010*) с использованием библиотеки OpenCV разработано программное обеспечение.

5.1. Распознавание изображений лиц

Целью исследования является поиск изображений лиц, которые соответствуют заданным тестовым образам. Численные эксперименты выполнены на основе базы изображений лиц [10]. Примеры изображений представлены на рис. 8.

База содержит 366 изображений лиц различных людей, по 20 изображений каждого лица. При формировании базы данных размер изображений и условия съемки были одинаковы. Применялся 24 битный формат JPEG. В базе [10] содержатся снимки людей мужского и женского пола, разных национальностей и возрастов. В ней отражены изменения внешности человека: различные причёски, присутствие бороды и очков.

В процессе подготовки к экспериментам были созданы две обучающие выборки. Первая из них содержит 5 изображений каждого человека (всего $5 \times 366 = 1830$ изображений). Вторая – 10 изображений лиц каждого человека для обучения (всего $10 \times 366 = 3660$ изображений). Тестовая выборка состоит из 20-ти изображений лиц каждого человека (всего $20 \times 366 = 7320$ изображений). Используются два типа вейвлет-преобразования: Хаара и Добеши. Результаты экспериментов показаны в таблице.

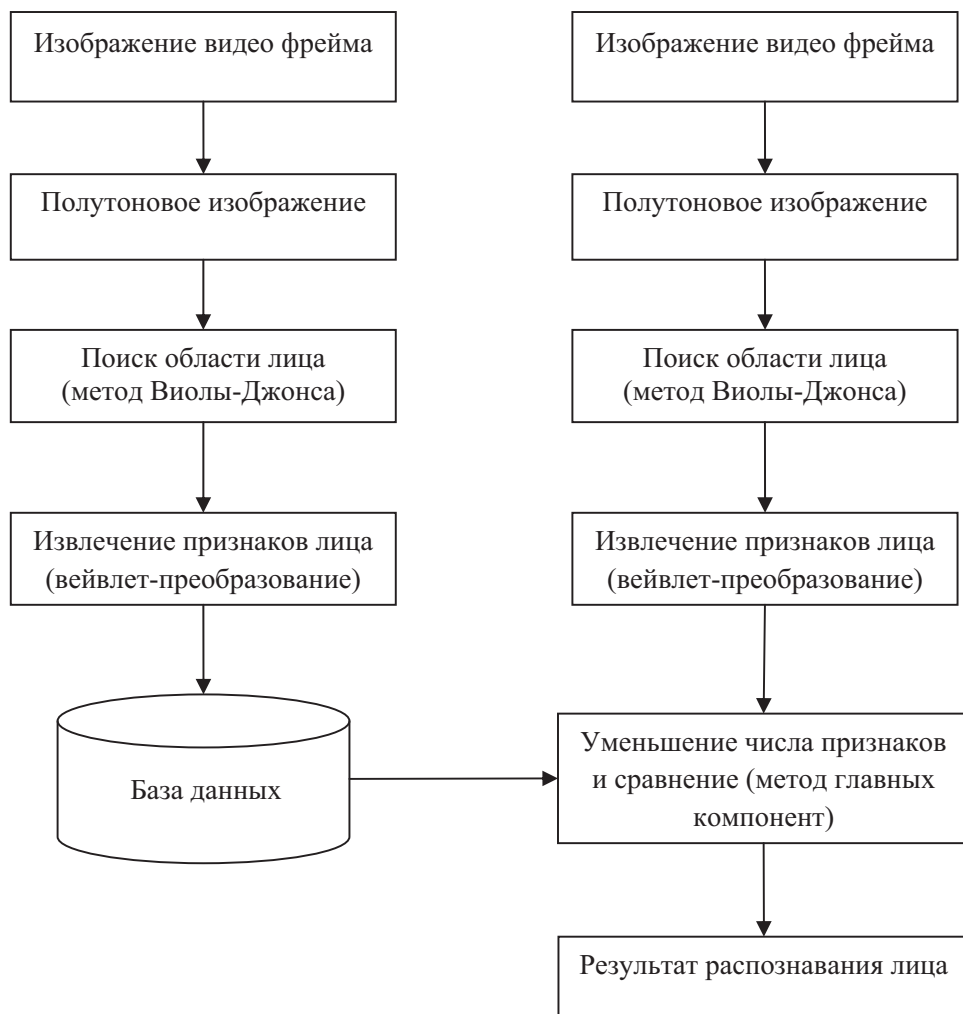


Рис. 7. Функциональная схема предложенного алгоритма распознавания лиц в видеопоследовательностях в режиме реального времени



Рис. 8. Примеры изображений лиц

Таблица. Количество распознаваний лиц

Тип вейвлет-преобразования	Хаар		Добеши	
	5 изоб.	10 изоб.	5 изоб.	10 изоб.
Достоверность распознаваний лиц, %	94,57	98,40	94,69	98,40

5.2. Распознавание лиц в видеопоследовательностях в режиме реального времени

На этапе обучения каждый человек находится перед камерой, совершая поступательные и вращательные движения головой. Программа записывает 10 различных изображений человека и сохраняет

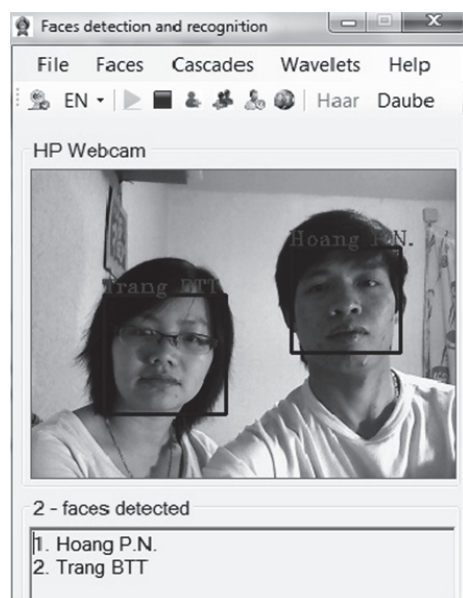


Рис. 9. Пример распознавания лиц в видеопоследовательностях в режиме реального времени

СПИСОК ЛИТЕРАТУРЫ

1. Viola P., Jones M.J. Rapid object detection using a boosted cascade of simple features // IEEE Conf. on Computer Vision and Pattern Recognition. – Kauai, Hawaii, USA, 2001. – V. 1. – P. 511–518.
2. Viola P., Jones M.J. Robust real-time face detection // International Journal of Computer Vision. – 2004. – V. 57. – № 2. – P. 137–154.
3. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.
4. Местецкий Л.М. Математические методы распознавания образов. – М.: МГУ, ВМиК, 2002–2004. – С. 42–44.
5. Sochman J., Matas J. AdaBoost with totally corrective updates for fast face detection // Center for Machine Perception, Czech Technical University, Prague. 2004. URL: <http://cmp.felk.cvut.cz/~sochmanj1/> (дата обращения: 10.02.2012).
6. Freund Y., Schapire R.E. A Short Introduction to Boosting // Journal of Japanese Society for Artificial Intelligence – September 1999. – V. 14. – № 5. – P. 771–780.
7. Буй Тхи Тху Чанг, Спицын В.Г. Разложение цифровых изображений с помощью двумерного дискретного вейвлет-преобра-

в базе данных с именем этого человека, рис. 9. Для проверки правильности работы программы этот процесс повторяется, и программа в реальном времени определяет – кто в данный момент находится перед видеокамерой. Скорость работы программы составляет 7–14 кадров/с.

Выводы

1. Предложен и описан новый алгоритм на основе совместного применения метода Виолы–Джонса, вейвлет-преобразования и метода главных компонент для распознавания лиц на изображениях и видеопоследовательностях в режиме реального времени.
2. Разработана программа, реализующая предложенный алгоритм распознавания лиц на изображениях и видеопоследовательностях в режиме реального времени на языке объектно-ориентированного программирования C#.
3. Анализ результатов компьютерных экспериментов позволяет говорить об успешной работе созданного алгоритма и программы при распознавании лиц.

- зования и быстрого преобразования // Известия Томского политехнического университета. – 2011. – Т. 318. – № 5. – С. 73–76.
8. Буй Тхи Тху Чанг, Фан Нгок Хоанг, Спицын В.Г. Алгоритмическое и программное обеспечение для классификации цифровых изображений с помощью вейвлет-преобразования Хаара и нейронных сетей // Известия Томского политехнического университета. – 2011. – Т. 319. – № 5. – С. 103–106.
 9. Брилюк Д., Старовойтов В. Распознавание человека по изображению лица и нейросетевые методы. 2002. URL: <http://daily.sec.ru/publication.cfm?pid=4425> (дата обращения: 10.02.2012).
 10. Spacek L. Description of the Collection of Facial Images. 2008. URL: <http://cswwww.essex.ac.uk/mv/allfaces/index.html> (дата обращения: 10.02.2012).

Поступила 24.02.2012 г.